PB-230 674

STOCHASTIC CONTROL AND MULTILEVEL COORDINATION MODELS FOR AIR QUALITY MANAGEMENT WITHIN A REGION

Ovadia A. Salama

Harvard University
Cambridge, Massachusetts

April 1971

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

PB 230 674

STOCHASTIC CONTROL AND MULTILEVEL COORDINATION MODELS FOR AIR QUALITY MANAGEMENT WITHIN A REGION

Ovadia A. Salama
Regional Science Group
Harvard University
April 1971

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

This paper was supported by grant #AP 00842-02 from the National Air Pollution Control Administration to the Regional Science Group, Graduate School of Design, Harvard University, Cambridge, Massachusetts.

TABLE OF CONTENTS

Page

TITLE PAGE.....
ABSTRACT.....
AKNOWLEDGMENTS.....
TABLE OF CONTENTS.....
LIST OF FIGURES.....

INTRODUCTION AND OVERVIEW.....

PART I: STOCHASTIC CONTROL MODELS OVER TIME FOR A SINGLE DECISION-MAKING UNIT.....

CHAPTER I: PROPOSED MODEL TO DETERMINE OPTIMAL CONTROL OF CONTAMINANTS EMISSIONS BY AN INDUSTRIAL PLANT.....

I.0 Context of the problem.....
I.I Role of the weather conditions.....
I.2 The abatement policy.....
I.3 The decision-making process.....
I.4 Problem formulation.....

CHAPTER 2: PROBLEM IDENTIFICATION.....

CHAPTER 3: ALGORITHM FOR RESOLUTION: CASE WHEN INTERPERIOD TRADE-OFFS ARE NOT PERMITTED.

3.I The principle of optimality.....
3.2 Numerical example.....
3.3 Computational results obtained with the numerical example.....

Page

CHAPTER 4: INCLUSION OF INTERPERIOD TRADE-OFFS IN THE MODEL.....

4.I Expression of the problem in the dynamic programming form.....

4.2 Application to the example developped in the previous section.....

CHAPTER 5: PROOF OF EXISTENCE AND UNICITY OF THE OPTIMAL SOLUTION BY THE ALGORITHM.....

5.I Alternative formulation of the problem.....

5.2 Solution of the problem.....

PART II: MULTILEVEL COORDINATION FOR SEVERAL DECISION-MAKING UNITS.....

CHAPTER 6: EXTENSION TO THE SITUATION WHERE SEVERAL PLANTS ARE PRESENT IN THE REGION.....

6.I Introduction of interactions between airsheds..

6.2 A model to take into account the interactions between airsheds.....

6.3 Analysis of the objective functions associated with the processes for a single time period....

6.4 Intra-level conflicts and the need for coordination.....

6.5 Interlevel conflicts.....

6.6 Modification of the taxation policy.....

6.7 The case when the variables are discrete.....

CHAPTER 7: ALGORITHM FOR THE STOCHASTIC CONTROL MODEL OVER TIME FOR SEVERAL COORDINATED DECISION-MAKING UNITS.....

7.I Summary of the procedure used by the single firm.....

Page

7.2 States of the system and cost functions in the multi-firm situation.....
7.3 Algorithm for the multi-period situation.....
7.4 Comparison with the solution of the overall problem.....

CONCLUSION: DIRECTIONS FOR FURTHER RESEARCH.....

APPENDIX.....
A.I Technology of waste gas treatment.....
A.2 Costs of waste gas treatment.....
A.3 Cost savings resulting from optimal operations..
A.4 Switching costs.....
a- start-ups.....
b- shut-downs and slow-downs.....

BIBLIOGRAPHY.....

"That second trap would be to let anti-pollution become our new multi-billion dollar business, to let the pollution go on merrily in all its present forms and superimpose a new multi-billion dollar business of anti pollution on top of it. And in these days of conglomerates it would be the same business; one branch of it would be polluting, the other would be anti-polluting".

George Wald

INTRODUCTION AND OVERVIEW

The present report is part of research aimed at examining how optimal control theory and hierarchical multilevel systems theory can be applied to some aspects of Air Quality problems within a region or city. The particular aspects we focus on are illustrated by the following statements of the assumed community goals, constraints, resources and objectives:

1) Overall goal. To keep the concentration of pollutants in the atmosphere during any time period lower than a given level whatever the meteorological conditions may be.

2) Overall constraint: To keep the output of the production sector the same as what it would be if this sector was not affected by the above requirement.

3) Basic resources: The technology to abate pollution is available and each potential polluter is already equipped (by law) with the adequate machinery capable to reduce its pollution emissions from no abatement at all to complete abatement. Between these two extremes, the machinery can be operated at various levels according to the needs.

The second basic resource is the limited and variable ability of the atmosphere to disperse pollutants. This ability is supposed to vary in a probabilistic way.

4) Basic costs: Actual use of the abatement machinery is associated with various costs, mainly costs of operating it at a given level and costs of switching from one level to another. (The cost of acquiring the abatement equipment is supposed to be integrated within the capital value of the plant in the same way emergency and fire protection equipment is imbedded within the construction cost of a building.)

On the other hand, social costs incurring to the community as a result of the presence of pollutants in the air are reflected by the more or less stringent character of the overall goal stated above, and are not analytically explicited otherwise.

5) Objectives: Each participant plant and the community as a whole seek over a set of time periods to reduce the cost of pollution abatement by using at best the two basic resources and satisfying the overall goal as closely as possible.

As stated above, the problem is one of finding the optimal control of a stochastic dynamic problem. According to the additional assumptions that can be made, this control must specify how the system consisting of all the polluting plants, should be set to operate at the beginning of every time period. We restrict ourselves here to memory-less processes. In other words, the changes in

weather conditions are assumed to follow a Markov process. This can be justified from the context of the problem.

As we are interested in the possibilities of practical applications and of actual computation of the controls, we limited ourselves to discrete-time, finite-set and finite horizon problems.

Introduction of Two Decisional Levels: The Control Problem

Our final aim is to study the Air Quality Management problem stated above at two decisional levels. The first level is that of the single production plant. The second level is that of the city as a whole, represented by a pooling or a monitoring agency.

With the single plant, we are dealing with a single-level, single-objective (though multi-variable) system. The study of this system as a first step of the analysis is necessary. It allows us to show on a relatively simple case the relevance of the problem, to introduce the terminology, the concepts and the partial models and to identify the variables of interest that are used in the more general model. Also, the solution algorithms for this small size problem can be used to pave the way for the study of larger systems and more complex cases.

Indeed, the aim of the pooling agency can be viewed as being to indicate to each plant, at the beginning of each time period, whether or not the plant should use its

emergency equipment during this period, given the goals, objectives, resources, constraints and costs introduced above. Again, we are dealing here with a single-level of decision making, but the size of the system -- measured in terms of the number of its possible state -- is much larger. Conceptually, the solution developed for the single plant problem can be transposed to solve the Agency's problem. Size considerations, however, suggested to investigate procedures for sharing the control responsibilities between the agency and the plants, which is the subject of the next paragraph.

Integration of the Two-Decisional Levels: the Coordination Problem

The set of plants located within the hypothetical area under study can be viewed as a single-level, multi-objective system, each participant plant having its own cost minimization objective function. If communication between participants is possible, compromise solutions leading to optimal use of the resources may be found through game theory and the theory of teams.

We have not been concerned with this particular approach. Rather, we chose to explore an alternative approach which implies integrating the individual plants and the pooling agency into a hierarchical, multi-level, decision-making system. In such a system, the agency represents the supremal level of decision making and has

a specific goal; the set of plants represents the infimal level of decision making and each plant has again a specific goal of a different type. Both of these goals are different from the goal of the community as stated in the first paragraph, which is the goal of the overall system. The search of the optimal use of the resources involves in that case a problem of coordination between the two levels of decision-making. A framework for studying this problem has been made available through the pioneering work of Walter Isard[1] and M.D. Mesarovic[2]. The application of these theories to the problem at hand is still under way and will be the subject of a forthcoming paper.

Objective of the Present Paper: A Single Plant Model,(Part 1)

The first part of this report is limited to the study of the case of the single-plant's decision problem. After describing and formulating the problem, we are concerned with, namely the optimal control of an industrial plant's emissions, taking into account weather changes, we propose a dynamic programming approach to solve it and we present an algorithm for the practical computations. We also present a formal proof of the existence and unicity of an optimal solution and proof of the fact that the pro-

---

[1]Walter Isard, General Theory, Chapter

[2]Mesarovic, et al., Theory of Hierarchical Multi-Level Systems, Academic Press, 1970.

posed algorithm actually leads to this optimal solution. The proof is based on the application of a version of Pontryagin's maximum principle to discrete problems. In addition, this proof also leads to the reformulation of the problem under a form which is better adapted to the generalizations to more complex systems presently under study.

In the second part, we indicate how this algorithm can be applied to the case of a central agency monitoring the use of the abatement equipment of all the plants in the region. Then, we examine how decision-making responsibility can be shared among two coordinated levels.

Part 1.

Stochastic Control Models Over Time for a Single Decision-Making Unit

The materials aimed at justifying the relevance of the problem are presented in a separate paper and are summarized in Appendix to the present one. These materials consist of data on various types of abatement equipment, the way it operates, its costs, etc. The main result pertinent to this part is that operating costs of the abatement equipment are of the same order of magnitude as the annual investment costs and that start-up and shut-down costs though variable, are not negligible and might affect the optimal use of this equipment.

Chapter 1. Proposed Model to Determine Optimal Control of Contaminants Emissions by an Industrial Plant

1.0. Context of the Problem

In a given area -- which is, for example, an industrial park near a large city -- the local atmosphere is considered as a geometric volume receiving pollutants emitted by a source located in the area. This source of pollutants is an industrial plant; for instance, an electric power generating plant.

We assume that this plant can use a technology chosen in a set of finite elements, say three elements [I,J,K]. The different technologies in question here involve using coals or fuels of different sulfur content, for example.

The can also refer to "situations where it is possible for the plant to schedule auxiliary operations, for which continuous production is not required, for periods of favorable weather

-or to operate collectors during periods of poor atmoshperic diffusion but not when diffusion is good, in order to save power, cost of scrubbing medium, or disposal cost of material collected

-or to schedule collection equipment shutdowns for regions for periods of good diffusion so that operations may be continued."[o]

---

[o] Air Pollution Handbook, P. Magill et al., McGraw-Hill, 1960.

Therefore, the word "technology" is used in the following in the meaning of level of use of the abatement equipment.

Each technology is associated with an operating cost and the set of possible technologies is completely ordered from the cheapest and dirtiest technology to the cleanest and most expensive one.

At a given period t (one day, for instance), the plant schedules the production of an output v_t. This can be done by using any of the technologies of the above set. We assume that at the beginning of a period it is possible to switch from the currently used technology to any other of the set. This can be done only once per period and at the beginning of the period. However, a switching cost is associated with the modification of the process from any technology of the set to another. These switching costs are summarized in a cost matrix M_t, where the entries of the main diagonal are zeroes.

At each time period a certain rate of pollutants, say of sulfur dioxide, is emitted in the local airshed. These pollutants are wiped away or carried over from one period to the next one according to the weather conditions.

1.1 Role of the Weather Conditions

The weather conditions which are relevant to air pollutant concentration[1] can be classified into a finite

[1] These can be combinations of mixing depths and wind speed levels.

number of types, say [E,F,G,...] and at the beginning of each time period the plant managers know with certainty from weather announcements what will be the type of weather for this time period. Moreover, we assume that there exists a transition matrix indicating the probability of type transition from one day to the next day.[2]

Let us assume that in the simplest case, the output level is fixed and remains the same over all the time periods. The only decision variables are then the technologies I,J,K. Then, these variables can be interpreted as emission quantities or emission rates as well as technologies.

The decision to choose one technology or another requires therefore examining the operating and switching costs, observing the weather conditions, and complying with the abatement policy enforced by the administrative power.

This problem is known in the air pollution control vocabulary as "meteorological management of air pollution". We shall now briefly discuss this latter point.

---

[2]The case where the probabilities of weather conditions depend on the weather of the 2 previous days can be treated similarly.

1.2 The Abatement Policy

At this stage, we wish to investigate the abatement policy that can be stated as follows: at any time period, the concentrations of pollutants in the local airshed should never exceed a given upper limit C.

A possible decision strategy of the plant that complies with such a condition can be obtained by calculating the concentrations resulting from the operations of any possible technology I,J,K,... (ordered from dirtiest to cleanest), in association with any type of weather condition E,F,G,... and by allowing to use only the combinations of technologies and weather types such that concentrations exceeding C can be wiped away during one period of time.

Several diffusion models are available for this computation.[3] Let us assume that by applying one of these we found that:

-weather type E allows the use of all I,J,K,...

-weather type F allows the use of J,K,...

-weather type G allows the use of K,...

-etc.

This means that if during one period technology I is used when the type of weather is E, all concentrations

---

[3] See, for example, Holzworth and Miller's model which is described in Appendix 3 of the previous paper.

of pollutants exceeding C can be wiped away during this period without any quantity of pollutants carried over to the next time period. In that case, each period of time is independent from the previous one as there is a natural cleaning mechanism at the beginning of each time period.

However, by doing so we do not take into acoount the fact that if, while weather conditions are of type E, we use technology J instead of technology I (which is the dirtiest one allowed with E) this would leave a margin of cleanliness in the air that should allow the use of a slightly dirtier technology during the following time period than the dirtiest technology allowed. If these inter-period trade-offs are possible, then, using J instead of I if the weather is E should make it somehow possible to use J in the following period instead of K if the weather conditions are of type G (the most constraining) and conversely.

In the following we shall state the problem for the general case where inter-period trade-offs are allowed. We shall then present solutions first for the special case when such trade-offs are not advisable, and secondly, for the general case.

1.3 The Decision Making Process

As far as the plant is concerned, we shall assume that its output is the same during all the time periods

and known. The market price of this output cannot be influenced by the plant. The switching and operating costs are given. The weather transition matrix is given. In these conditions, the profit maximizing behaviour of the firm implies minimizing the total cost over the time horizon of the firm's operations.

We shall study the case where the time horizon is finite. Costs can be either discounted or not.

At the beginning of time period t, the plant is operating with a certain technology, say I. The plant management will examine the weather announcement for this period. This will show that his process can be continued with I or has to be switched down, say one level down, i.e. that he has to switch to J. However, he may choose to switch to a lower level, that is, K for instance, if he expects that in the following period (t+1) there is a large enough chance that the weather type will be G. By doing so, he would operate with K during period t, incurring a higher cost than if he operated with J, but avoiding two successive switching costs if his expectation for the weather in period t+1 turns out to be true. Our problem is to define how the manager will choose each step so that he minimizes his total cost over the horizon, to determine what is this minimal cost and how much this cost differs from the minimal cost without abatement strategy which is clearly to use the cheapest and dirtiest technology, I, all the time.

1.4 Problem Formulation

It will be convenient to consider the system formed by the plant and its environment and to define the states $S = \{1,\ldots,s\}$, of this system as couples (technology-weather type). In the example above, the number of possible states of the process is 9, and this set is the following:

$$S = \{IE, IF, IG, JE, JF, JG, KE, KF, KG = 1,2,3,\ldots 9\}$$

We shall number these states so that we can in the following denote a state by its number.

We suppose that the state of the process is observed at the beginning of each period $t = 0,1,\ldots,T-1$. Let $x_t \in S$ denote the state observed at the beginning of period t. After observation one of a possible number of decisions is made, namely to stay in the same state x_t or to switch to another state $\bar{x}_t$ belonging to a subset $R(x_t) \in S$.

Let $\{x_t\}$ and $\{\bar{x}_t\}$, $t = 0,1,\ldots T-1$ be the sequence of states of the system before and after a decision is made and let $\{\Delta_t\}$, $t = 0,1,\ldots,T-1$ be the sequence of decisions.

It is possible to associate with each decision Δ_t a number u_t such that we can write $\bar{x}_t = x_t + u_t$. This number u_t can be zero, positive or negative, provided

$x_t \in S$, that is, for the example above $1 \leq x_t \leq 9$. Let $U(\bar{x}_t)$ be the set of possible u_t associated with $\bar{x}_t$.

After a decision is made and the process placed in state $\bar{x}_t = x_t + u_t$, an exogenous-random change takes place to move the process from $\bar{x}_t$ to another state, called x_{t+1} which can be observed at the beginning of time period (t+1).

We assume that this is done according to a Markov process whose matrix of transitions from one time period to the next is $\phi_t(i,j)$, $i, j \in S$. We assume that the transition probabilities do not depend on past history once a decision to switch or to remain in a given state has been made. By definition:

$$\phi_t(i,j) = \Pr \{x_{t+1} = j \mid \bar{x}_t = i\}$$

Some entries of this matrix are equal to zero, namely, those which refer to transitions from one state to another involving switching technologies, which cannot be done spontaneously.

There are costs associated with being in a certain state. They are the operating costs. If it is not allowed to be in a given state, the corresponding operating cost is equal to infinity. Let $L_t(x_t)$ be these costs.

Switching from one state to another is also associated with a cost designated by $M(x_t, \bar{x}_t)$. As it is not

possible to switch from one weather type to another, the corresponding switching costs from one weather type to another will also be equal to infinity.

The total cost incurred at each period is then:

$$c_t = L_t(\bar{x}_t) + M_t(x_t, \bar{x}_t).$$

We introduce now a constraint on the pollutants concentration buildup. Let y_t be the change in the concentration of pollutants at time t resulting from the operations of the plant in state $\bar{x}_t$. We assume that it is possible to measure the value of y for each state of the system by means of a diffusion model. For instance, let us assume that in our example, we found that the set y, defined as a vector, each component of which is associated with one of the elements of the set S, is the following:

state	y
IE	0
IF	α
IG	2α
JE	$-\alpha$
JF	0
JG	$+\alpha$
KE	-2α
KF	$-\alpha$
KG	0

where α is a unit of concentration of pollutants in the atmosphere.

The constraints on the build-up of pollutants concentration in the atmosphere can then be written as:

$$\sum_{\tau=0}^{t} y_\tau \leq C \qquad \text{for all } t$$

$$\sum_{\tau=0}^{t} y_\tau \geq 0 \qquad \text{for all } t$$

The problem of interest is to find the sequence of $\{u_t\}$ that minimizes the total expected cost of operations and switchings within the planning horizon, without letting the concentration of pollutants of the atmosphere at period t, namely $\sum_t y_t$, exceed a certain fixed quantity C.

In formal terms:

$$\min_{\{u_t\}} \quad E\{\sum_t L_t(x_t + u_t) + M_t(x_t, x_t + u_t)\}$$

subject to

$$\sum_t y_t \leq C \qquad \text{for all } t$$

$$\sum_t y_t \geq 0 \qquad \text{for all } t$$

$$x_t + u_t > 0 \qquad \text{for all } t$$

$$x_t > 0 \qquad \text{for all } t$$

such that $\phi_t(i,j) = \Pr\{x_{t+1} = j | \bar{x}_t = i\}$ is a given transition matrix.

In order to make more explicit the dynamics of the system, we can define the random variable $r_t = x_{t+1} - \bar{x}_t$ whose probability distribution can be derived from $\phi_t(i,j)$ and write

$$x_{t+1} = x_t + u_t + r_t \quad , \quad x_o \ S$$

where x_t and x_{t+1} are the state of the system observed at the beginning of time periods t and t+1, u_t is the decision variable chosen during time period t and r_t the random variable responsible for the spontaneous change in the state of the system during time period t.

Finally, in order to show the kind of relationship linking y_t with the objective function, let us introduce a vector δ_K whose dimension is equal to the number of elements of the set of states S and whose components are all equal to zero except for the Kth which is equal to one:

$$\delta_K = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \quad \text{Kth component}$$

Let δ_K' denote the transpose of δ_K.

Our problem can be restated in the following terms, where L_t is the vector of operating costs associated with the elements of the set S.

$$\min_{u} \quad E\{\Sigma_t \delta'_{x_t+u_t} \cdot L_t + \delta'_{x_t} \cdot M_t \cdot \delta_{u_t}\}$$

subject to $\quad \Sigma_t \delta'_{x_t+u_t} \cdot y \leq C$

$$\Sigma_t \delta'_{x_t+u_t} \cdot y \geq 0 \qquad \text{for all } t$$

$$x_t + u_t > 0$$

$$x_t > 0$$

where $x_{t+1} = x_t + u_t + r_t$

and $\phi_t(i,j) = Pr\{x_{t+1} = j | x_t + u_t = i\}$ is a given transition matrix.

The nature of the problem depends essentially on how we define a solution. In order to do so, let us consider the matrix D_t whose entries $D_t(i,j)$ are defined as follows:

$$D_t(i,j) = \text{Prob}\,[u_t + i | x_t = j]$$

$$0 \leq D_t(i,j) \leq 1 \qquad \forall\, i,j,t$$

$$\sum_{i=1}^{g} D_t(i,j) = 1 \qquad \forall\, j,t$$

The variables $D_t(i,j)$ will play the role of control. The matrix D_t will be called "decision rule" or "control rule" or "strategy".

The value of u_t that has to be used at each instant t is sought so as to minimize the objective function. In otherwords, the optimal value of the matrix of probabilities D_t has to be determined. If the optimal value of these variables is either 1 or 0, then the choice of the input is unambiguous. The strategy is called pure or purely deterministic stationary. If it is not possible to find a pure strategy, then a mixed (randomized) strategy has to be sought.

It is possible that no optimal rule D_t will exist. It is also possible that an optimal rule may exist but not be of the purely deterministic stationary type.

Chapter 2. Problem Identification

The problem presented above belongs to the class of discrete optimal control problems. A discrete optimal control problem can be described in a fairly general way. This problem is concerned with maximizing some objective function over some finite planning horizon, beginning with a fixed initial state and a set of difference equations which describe the dynamics of the system. There are many other things which could be considered, such as

constraints on the state variables (both at the terminus and/or along the trajectory), constraints linking the feasible control region of one period to the state variables of the previous period, objective functions whose values depend only on the terminal state, etc. However, such issues need not be added to this part of the discussion. They can easily be included once the main thrust has been made.

In particular, let us consider a general system whose state at the beginning of any period t, can be described by the n-vector x_t. The initial state will be denoted x_o. During each period the control (or decision) variables which may be applied are represented by the r-vector u_t. The interaction of this control acting upon a system at x_t determines the movement of the system and is represented by the set of simultaneous difference equations

$$\Delta x_t = d(x_t, u_t)$$

where Δx_t is the difference operator $x_{t+1}-x_t$. Usually the control vector is constrained to be selected for a feasible region defined by a set of inequalities

$$U = [u_t | g_i(u_t) \leq 0, \quad i = 1,\ldots,k]$$

Finally, the objective function of each period by itself depends both on the state of the system and on the control input and is denoted by $\mu_t(x_t, u_t)$. Thus for a

planning horizon of T periods, the whole optimal control problem is described as that of

$$\text{minimizing} \quad \sum_{t=0}^{T-1} \mu_t(x_t, u_t)$$

$$\text{subject to} \quad \Delta x_t = d(x_t, u_t) \qquad t = 1, \ldots, T-1$$

$$u_t \in U \quad , \quad x_o \text{ given}$$

This scheme is more complicated when we deal with a system where random events can occur. The formulation depends on whether the random events occur before or after the decision is made. In any case, the criterion is replaced by the expected value of the above objective function and the dynamics of the system reflect in the appropriate way the occurence of the random events.

The correspondences with our problem are quite clear:

i) x_t is a vector of variables which characterizes the state of the system at the beginning of each time period. It is a 3-dimensional vector which has as components the weather, and the pollution concentration.

ii) u_t is the vector of variables which characterizes the decision to use a given state during the period, state where the technology and the pollution concentration are changed.

iii) U is defined by the set of constraints expressing the fact that the concentration of pollutants should

not exceed a given level during any period.

iv) $d(x_t,u_t)$ is the change in the state of the system caused by the random change of the weather conditions. It is defined by the two following equations:

$$x_{t+1} = x_t + u_t + r_t$$

$$\phi_t(i,j) = Pr\{x_{t+1} = j | x_t + u_t = i\}$$

v) $\mu_t(x_t,u_t)$ is equivalent to the sum of $L_t(u_t)$ + $M_t(x_t,u_t)$ in our model. Actually, we consider the expected value of this quantity which represents the total expected cost of operating the abatement equipment.

Chapter 3. Algorithm for Resolution: Case when Interiod Trade-offs are not Permissible

In that case a certain number of simplifications can be introduced in the general model. Here, the concentration of pollutants at the end of each period have to be equal to zero; that is, the only technologies allowed are those whose associated emissions are either equal or less than the capacity of the atmosphere to eliminate them during one time period. Therefore, the only states allowed at the beginning of each period are those for which y is either zero or negative. The restricted set of possible decisions is then:

$u_t \in U = \{IE, JE, JF, KE, KF, KG\}$.

This set is immediately available from the application of the diffusion models. It is the same for all time periods and in particular it is independent from the state of the system in the previous time period. The constraints have no longer any use and can be discarded. The problem is reduced to the following one: choose at every time period the control $u_t \in U$ which minimizes the objective function $\sum_t E[L_t(u_t) + M_t(x_t,u_t)]$ where x_t belongs to the set S of reachable states $S = \{IE,IF,IG,JE,JF,JG,KE,KF,KG\}$. Knowing that, after a decision u_t is made, the system will evolve under random influence towards a state x_{t+1} depending only on u_t, such as the probabilities

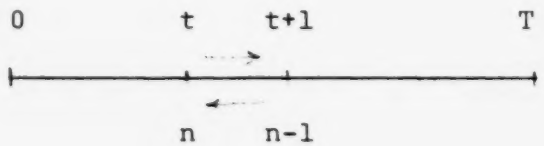
$$\phi_t(i,j) = \Pr\{x_{t+1} = j \mid \bar{x}_t = i\}$$

are known.

The solution to this simplified problem is obtained by applying the dynamic programming algorithm as indicated in the following paragraphs.

3.1 The Principle of Optimality (Dynamic Programming)

Let t be the forward-running time index and n be the backward-running stages-to-go index. We assume that the phenomenon is expanding on T periods of time



Using $f_n(x_t)$ to denote the optimal value attainable when the current state is x_t and there are n stages to go, we have the general statement of the principle of optimality

$$f_n(x_t) = \min_{u_t \in U} [L_t(u_t) + M_t(x_t,u_t) + f_{n-1}(x_{t+1})]$$

and we are ultimately interested in finding $f_t(x_o)$ and the u_t^*-sequence of optimal controls that produces it.

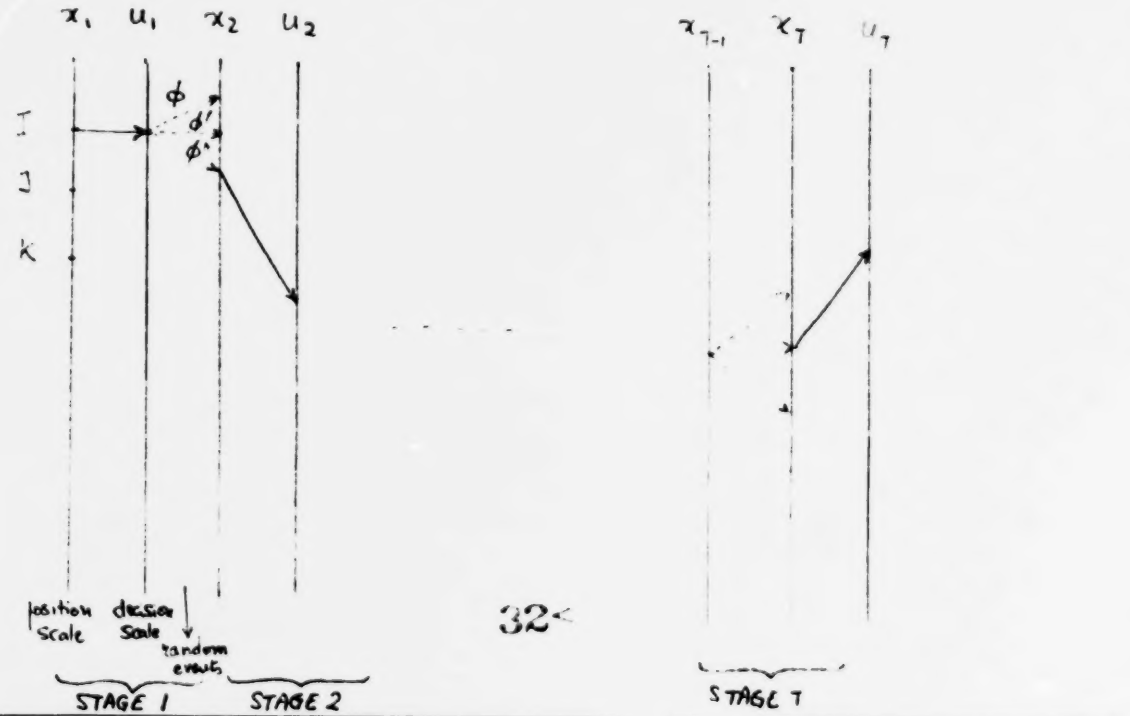
The above equality expresses that an optimal substrategy from T-t to T can only be made of optimal substrategies from T-t+1 to T and this applies for all t from 0 to T. In other words, an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

This principle applied to stochastic situations requires the use of the expected value criterion It implies looking for the optimum by backing up from the future to the past as n is the backward-running stages-to-go index. In that case, $f_n(x_t)$ has to be interpreted as the expected optimal value attainable when the current state is x_t and there are n stages to go.

A discount factor β_t can be introduced for time period t. The discounted cost for period t is the product $\beta,x...x\beta_t$ times the cost in period t. In case all $\beta_t=\beta$, the discounted cost is just β^t times the cost. However,

we shall not explicitly introduce this factor. Its presence in the objective function can be noted by the presence of the subscript t attached to the costs L and M.

Dynamic programming problems can be represented diagrammatically as follows. For each stage the possible states x_t of a system are represented by points on a scale called the position scale and the possible decisions u_t are represented by points on a parallel scale called the decision scale. A decision to move from a point on the position scale to a point on the decision scale is represented by a bundle of dotted arrows, each one associated with its probability of occurence. In the case when the random events occur after the decision is made, these arrows join points of the decision scale of one stage to points of the position scale of the next stage. This type of diagram is generally used to summarize the information needed to solve the problem.



3.2 Numerical Example

Let us assume that the probabilities of weather types transitions from one day to the other are given by the entries of the following matrix:

	E	F	G
E	.5	.4	.1
F	.3	.2	.5
G	.2	.4	.4

Operating costs for the allowed combinations technology-weather (that is, associated with a zero concentration of pollutants at the end of the time period) are given by the following vector:

$$L(u) = \begin{matrix} L(IE) \\ L(JE) \\ L(JF) \\ L(KE) \\ L(KF) \\ L(KG) \end{matrix} = \begin{matrix} 10 \\ 20 \\ 20 \\ 30 \\ 30 \\ 30 \end{matrix}$$

The matrix of switching costs is the following:

	I	J	K
I	0	20	10
J	20	0	30
K	10	30	0

and the total cost matrix (switching plus operating costs) is therefore:

	I	J	K
I	10	40	40
J	30	20	60
K	20	50	30

These matrices can be rewritten as follows:

Matrix of State Transitions: (all non indicated entries are zero)

$\phi_t(i,j)$ =

	IE	IF	IG	JE	JF	JG	KE	KF	KG
IE	.5	.4	.1						
IF	.3	.2	.5						
IG	.2	.4	.4						
JE				.5	.6	.1			
JF				.3	.2	.5			
JG				.2	.4	.4			
KE							.5	.4	.1
KF							.3	.2	.4
KG							.2	.4	.4

Switching Costs Matrix: (all missing entries correspond to states which are not allowed, i.e., which corresponding costs are infinite).

$M_t(x_t, \bar{x}_t)$ =

	IE	IF	IG	JE	JF	JG	KE	KF	KG
IE	10	.	.	40	.	.	40	.	.
IF	.	10	.	.	.	.	.	.	.
IG	.	.	10	.	.	.	.	.	.
JE	30	.	.	20	.	.	60	.	.
JF	.	30	.	.	20	.	.	60	.
JG	.	.	.	.	.	20	.	.	.
KE	20	.	.	50	.	.	30	.	.
KF	.	20	.	.	50	.	.	30	.
KG	.	.	20	.	.	50	.	.	30

The following figure summarizes the information on this problem:

1st Stage				2nd Stage	
position scale	decision scale		random event		
		.5	IE		
IE	IE	.4	IF	IE	IE
		.4	IG		
	JE	.5	JE		JE
		.4	JF		
		.1	JG		
	KE	.5	KE		KE
		.4	KF		
		.1	KG		
IF	JF	.3	JE	IF	JF
		.2	JF		
		.5	JG		
	KF	.3	KE		KF
		.2	KF		
		.5	FG		
IG	KG	.2	KE	IG	KG
		.4	KF		
		.4	KG		
JE	IE	.5	IE	JE	IE
		.4	IF		
		.1	IG		
	JE	.5	JE		JE
		.4	JF		
		.1	JG		
	KE	.5	KE		KE
		.4	KF		
		.1	KG		
JF	JF	.3	JE	JF	JF
		.2	JF		
		.5	JG		
	KF	.3	KE		KF
		.2	KF		
		.5	KG		
JG	KG	.2	KE	JG	KG
		.4	KF		
		.4	KG		

KE	IE	.5	IE	KE	IE
		.4	IF		
		.1	IG		
	JE	.5	JE		JE
		.4	JF		
		.1	JG		
	KE	.5	KE		KE
		.4	KF		
		.1	KG		
KF	JF	.3	JE	KF	JK
		.2	JF		
		.5	JG		
	KF	.3	KE		KF
		.2	KF		
		.5	KG		
KG	KG	.2	KE	KG	KG
		.4	KF		
		.4	KG		

Let the planning horizon be T periods. That is, we consider the periods beginning at times 0,1,...T-1. In period t, $0 \leq t < T-1$, it is convenient to focus on the number $n = T-t$ of remaining periods.

If the system is in state $x_t = i \in S$, according to the principle of optimality, the decision to stay in i or to switch to $r \in U(x_t)$ is made as follows.

Let $f_n(x_t)$ denote as indicated above, the minimum total expected cost for the n periods t, t+1,...,T-1.

Let $\phi_t(S,S)$ represent the complete matrix whose elements are $\phi_t(i,j)$, $\forall i,j$. For any subset M and K of S, $\phi_t(M,K)$ will represent the submatrix of $\phi_t(S,S)$ formed by taking rows $i \in M$ and columns $j \in K$.

In particular, $\phi_t(i,S) = \{\phi_t(i,1),\ldots,\phi_t(i,S)\}$. Also, if f is a function of S and M any subset of S, f(M) will represent the vector

$$\begin{matrix} f_{(m_1)} \\ f_{(m_2)} \\ \\ f_{(m_n)} \end{matrix} \qquad \text{where } M = \{m_1 \ldots, m_n\}$$

The values $f_n(i)$ are first determined inductively by

$$f_n(i) = \min \left[L_t(i) + \phi(i,S)\, f_{n-1}(S), \right.$$

$$\left. \min_{r \in U} M_t(i,r) + L_t(r) + \phi(r,S)\, f_{n-1}(s) \right\} \; ; \; i \in S.$$

Let $L_t(i) + \phi(i,S)\, f_{n-1}(S) = g_n(i)$; $i \notin S$

Then $f_n(i) = \min [g_n(i), \min_{r \in U} \{M_t(i,r) + g_n(r)\}]$

Let $r_n \in U$ be the value that minimizes $M_t(i,r)+g_n(r)$,

that is: $M_t(i,r_n) + g_n(r_n) = \min_{r \in U} [M_t(i,r) + g_n(r)]$

and let $M_n = [i \in S;\ g_n(i) \leq M_t(i,r_m) + g_n(r_n)$

and $K_n = [i \in S,\ i \notin M_n]$

Then $f_n(i) = g_n(i)$ if $i \in M_n$

$f_n(i) = M_t(i,r_n) + g_n(r_n)$ if $i \in K_n$.

The optimal policy specified by the pair (M_n,r_n) is to switch in period t = T-n if and only if the state is not in M_n, that is, remain in state i if i M_n and switch to state r_n if i K_n.

In order to apply this iteration technique to our numerical example, let $z_n(x_t,u_t)$ for $u_t \in U$ be the expected cost when the system is in x_t and is moved to u_t.

$$f_n(x_t) = \min z_n(x_t,u_t)$$

In our example, the last stage (time period T-1, stages to go n = 1) is deterministic as the problem ends

before random events occur. The optimal strategy is then easy to calculate as, for that stage the functions

$z_n(x_t, u_t)$ are the costs (as opposed to the expected costs) of all combination pairs of state and decision variables. The best move corresponds to the minimum of z_n.

This is summarized in the following table (Table 1) where the optimal decision u_{T-1} are indicated, corresponding to all possible states x_{T-1} where the system can be found at the beginning of the last period.

TABLE 1

Last Period: Stages to go N=1 Position Variable x_{T-1}	Decision Variables Allowed u_{T-1}	Expected Costs $z_2(x_{T-1}, u_{T-1})$	Best Decision u^*_{T-1}	Corresponding Cost $f_2(x_{T-1}) = \min z_2(x_{T-1}, u_{T-1})$
IE	IE	10	stay in I	10
	JE	40		
	KE	40		
IF	JF	40	indifferent between J and K	40
	KF	40		
IG	KG	40	switch to K	40
JE	IE	30	stay in J	20
	JE	20		
	KE	60		
JF	JF	20	stay in J	20
	KF	60		
JG	KG	60	switch to K	60
KE	IE	20	switch to I	20
	JE	50		
	KE	30		
KF	JF	50	stay in K	30
	KF	30		
KG	KG	30	switch to K	30

For stage t = T-2 (stages to go n = 2), we also have to examine all possible position variables x_{T-2}. If the system is in x_{T-2} at the beginning of time period T-2, it can be switched to one of the positions $u_{T-2} \in \{U\}$. The set $\{U\}$ is in the present case the same for all time periods. It does not depend on previous or following values of x_t or u_t. This is what makes the problem simpler than the general case stated at the outset.

After the choice of a state u_{T-2}, the system will be moved at random to x_{T-1}.

The optimal expected cost over the last two periods corresponding to each positon variable x_{T-2} is computed by combining x_{T-2} with each of the elements u_{T-2} of the decision set and by following it through period (T-1) where the system can be switched at random to one out of three possible new positions.

The total cost will be the sum of the cost of operating in period T-2 (operating and switching costs) and the expected minimum cost to be incurred during period T-1.

This procedure is indicated in the following table (Table 2) which indicates for each x_{T-2} what is the best move u^*_{T-2} given the uncertainties expected in the following stage.

TABLE 2

Period t = T-2 (Stages to go n = 2)

Position Variable x_t	Decision Variables Allowed u_t	Expected Costs $z_n(x_t,u_t)$	Best Decision u_t^*	Corresponding Cost $f_n(x_t)=\min z_n(x_t,u_t)$
IE	IE	10 + .5(10) + .4(40) + .1(40)= 35	STAY IN I	35
	JE	40 + .5(20) + .4(20) + .1(60)= 64		
	KE	40 + .5(20) + .4(30) + .1(30)= 65		
IF	JF	40 + .3(20) + .2(20) + .5(60)= 80		
	KF	40 + .3(20) + .2(30) + .5(30)= 72	SWITCH TO K	72
IG	KG	40 + .2(20) + .4(30) + .4(30)= 68	SWITCH TO K	68
JE	IE	30 + .5(10) + .4(60) + .1(40)= 55		
	JE	20 + .5(20) + .4(20) + .1(60)= 44	STAY IN J	44
	KE	60 + .5(20) + .4(30) + .1(30)= 85		
JF	JF	20 + .3(20) + .2(20) + .5(60)= 60	STAY IN J	60
	KF	60 + .3(20) + .2(30) + .5(30)= 92	SWITCH TO K	88
JG	KG	60 + .2(20) + .4(30) + .4(30)= 88		
KE	IE	20 + .5(10) + .4(40) + .1(40)= 45	SWITCH TO I	45
	JE	50 + .5(20) + .4(20) + .1(60)= 74		
	KE	30 + .5(20) + .4(30) + .1(30)= 55		
KF	JF	50 + .3(20) + .2(20) + .5(60)= 90		
	KF	30 + .3(20) + .2(30) + .5(30)= 62	STAY IN K	62
KG	KG	30 + .2(20) + .4(30) + .4(30)= 98	STAY IN K	98

Eventually, for any time period t (stages to go n+T-t) it is possible to derive the best move associated with each of the nine positions where the system can be found at the beginning of this time period, provided we know the nine optimal values of the function $f_{n-1}(x_{t+1})$ associated with the remaining stages.

The algorithm consists, as shown in Table 3, of feeding at the time period t the values of $f_{n-1}(x_{t+1})$ in the following matrix.

$$M = \begin{bmatrix} .5(\quad) + .4\ (\quad) + .1\ (\quad) \\ .5(\quad) + .4\ (\quad) + .1\ (\quad) \\ .5(\quad) + .4\ (\quad) + .1\ (\quad) \\ .3(\quad) + .3\ (\quad) + .5\ (\quad) \\ .3(\quad) + .3\ (\quad) + .5\ (\quad) \\ .2(\quad) + .2\ (\quad) + .4\ (\quad) \end{bmatrix}$$

by multiplying this matrix by the matrix (f_{n-1}) defined as follows:

$$[f_{n-1}] = \begin{bmatrix} f_{n-1}(IE) & f_{n-1}(IE) & f_{n-1}(IE) & f_{n-1}(JE) & f_{n-1}(JE) & f_{n-1}(KE) \\ f_{n-1}(IF) & f_{n-1}(IF) & f_{n-1}(IF) & f_{n-1}(JF) & f_{n-1}(JF) & f_{n-1}(KF) \\ f_{n-1}(IG) & f_{n-1}(IG) & f_{n-1}(IG) & f_{n-1}(JG) & f_{n-1}(JG) & f_{n-1}(KG) \end{bmatrix}$$

Then adding the vector

$$P(I) = \begin{bmatrix} 10 \\ 40 \\ 40 \\ 40 \\ 40 \\ 40 \end{bmatrix} ;$$

taking the minimum corresponding to the first 3 rows, calling it $f_n(IE)$;

taking the minimum of the 4th and 5th row and calling it $f_n(IF)$;

taking the minimum of the last term and calling it $f_n(IG)$.

Then, doing the same operation with vectors $P(J) = \begin{matrix} 30 \\ 20 \\ 60 \\ 20 \\ 60 \\ 60 \end{matrix}$

and $P(K) = \begin{matrix} 20 \\ 50 \\ 30 \\ 50 \\ 30 \\ 30 \end{matrix}$

which gives us respectively the values of $f_n(JE)$, $f_n(JF)$, $f_n(JG)$ and $f_n(KE)$, $f_n(KF)$, $f_n(KG)$.

that complete the computation of the nine values of the function $f_n(x_t)$.

These nine values should be rearranged under the following matrix form for the convenience of the next step's calculations

$$[f_n] = \begin{bmatrix} f_n(IE) & f_n(IE) & f_n(IE) & f_n(JE) & f_n(JE) & f_n(KE) \\ f_n(IF) & f_n(IF) & f_n(IF) & f_n(JF) & f_n(JF) & f_n(KF) \\ f_n(JG) & f_n(IG) & f_n(IG) & f_n(JG) & f_n(JG) & f_n(KG) \end{bmatrix}$$

This process is repeated until t = 0 which completes the definition of the optimal solution associated with each starting point.

A computer program can easily be written to undertake these calculations.

3.3 Computational Results Obtained with the Numerical Example

The computations have been carried on manually for the first 6 iterations, and are presented in Tables 3 through 7.

A close examination of these results shows that for most of the cases the strategy is stable, that is at every stage the same decision is associated with the same position variable. At the 6th iteration one change occurs: when the starting technology is J and the weather conditions are E, the optimal strategy becomes "switch to I" instead of "stay in J". The figures show that this strategy is stable when the process involves even more stages: it is a steady-state strategy.

TABLE 3

Period t = T-2 (Stages to go n = 2)

Position Variable x_t	Decision Variables Allowed u_t	Expected Costs $z_n(x_t,u_t)$	Best Decision u_t^*	Corresponding Cost $f_n(x_t)=\min z_n(x_t,u_t)$
IE	IE	10 + .5( ) + .4( ) + .1( )=		
	JE	40 + .5( ) + .4( ) + .1( )=		
	KE	40 + .5( ) + .4( ) + .1( )=		
IF	JF	40 + .3( ) + .2( ) + .5( )=		
	KF	40 + .3( ) + .2( ) + .5( )=		
1G	KG	40 + .2( ) + .4( ) + .4( )=		
JE	IE	30 + .5( ) + .4( ) + .1( )=		
	JE	20 + .5( ) + .4( ) + .1( )=		
	KE	60 + .5( ) + .4( ) + .1( )=		
JF	JF	20 + .3( ) + .2( ) + .5( )=		
	KF	60 + .3( ) + .2( ) + .5( )=		
JG	KG	60 + .2( ) + .4( ) + .4( )=		
KE	IE	20 + .5( ) + .4( ) + .1( )=		
	JE	50 + .5( ) + .4( ) + .1( )=		
	KE	30 + .5( ) + .4( ) + .1( )=		
KF	JF	50 + .3( ) + .2( ) + .5( )=		
	KF	30 + .3( ) + .2( ) + .5( )=		
KG	KG	30 + .2( ) + .4( ) + .4( )=		

TABLE 4

Period t = T- 3 (Stages to go n = 3)

Position Variable x_t	Decision Variables Allowed u_t	Expected Costs $z_n(x_t,u_t)$		Best Decision u_t^*	Corresponding Cost $f_n(x_t)=\min z_n(x_t,u_t)$
IE	IE	10 + .5(35) + .4(72) + .1(68)	= 63.1	STAY IN I	63.1
	JE	40 + .5(44) + .4(60) + .1(88)	= 94.8		
	KE	40 + .5(45) + .4(62) + .1(58)	= 93.1		
IF	JF	40 + .3(44) + .2(60) + .5(88)	= 109.2	SWITCH TO K	94.5
	KF	40 + .3(45) + .2(62) + .5(58)	= 94.5		
IG	KG	40 + .2(45) + .4(62) + .4(58)	= 97.0	SWITCH TO K	97.0
JE	IE	30 + .5(35) + .4(72) + .1(68)	= 83.1		
	JE	20 + .5(44) + .4(60) + .1(88)	= 74.8	STAY IN J	74.8
	KE	60 + .5(45) + .4(62) + .1(58)	= 103.1		
JF	JF	20 + .3(44) + .2(60) + .5(88)	= 89.2	STAY IN J	89.2
	KF	60 + .3(45) + .2(62) + .5(58)	= 114.5		
JG	KG	60 + .2(45) + .4(62) + .4(58)	= 117.0	SWITCH TO K	117.0
KE	IE	20 + .5(35) + .4(72) + .1(68)	= 73.1	SWITCH TO I	73.1
	JE	50 + .5(44) + .4(60) + .1(88)	= 104.8		
	KE	30 + .5(45) + .4(62) + .1(58)	= 83.1		
KF	JF	50 + .3(44) + .2(60) + .5(88)	= 119.2	STAY IN K	94.5
	KF	30 + .3(45) + .2(62) + .5(58)	= 84.5		
KG	KG	30 + .2(45) + .4(62) + .4(58)	= 87.0	STAY IN K	87.0

TABLE 5

Period t = T- 4(Stages to go n =4)

Position Variable x_t	Decision Variables Allowed u_t	Expected Costs $z_n(x_t,u_t)$		Best Decision u_t^*	Corresponding Cost $f_n(x_t)=\min z_n(x_t,u_t)$
IE	IE	10 + .5(63.1) + .4(94.5) + .1(97.0)	= 89.05	STAY IN I	89.05
	JE	40 + .5(74.8) + .4(89.2) + .1(117.0)	= 124.78		
	KE	40 + .5(73.1) + .4(84.5) + .1(87.0)	= 119.05		
IF	JF	40 + .3(74.8) + .2(89.2) + .5(117.0)	= 138.48	SWITCH TO K	122.33
	KF	40 + .3(73.1) + .2(84.5) + .5(87.0)	= 122.33		
IG	KG	40 + .2(73.1) + .4(84.5) + .4(87.0)	= 123.22	SWITCH TO K	123.22
JE	IE	30 + .5(63.1) + .4(94.5) + .1(97.0)	= 109.05		
	JE	20 + .5(74.8) + .4(89.2) + .1(117.0)	= 104.78	STAY IN J	104.78
	KE	60 + .5(73.1) + .4(84.5) + .1(87.0)	= 139.05		
JF	JF	20 + .3(74.8) + .2(89.2) + .5(117.0)	= 118.48	STAY IN J	118.48
	KF	60 + .3(73.1) + .2(84.5) + .5(87.0)	= 142.33		
JG	KG	60 + .2(73.1) + .4(84.5) + .4(87.0)	= 143.22	SWITCH TO K	143.22
KE	IE	20 + .5(63.1) + .4(94.5) + .1(97.0)	= 99.05	SWITCH TO I	99.05
	JE	50 + .5(74.8) + .4(89.2) + .1(117.0)	= 134.78		
	KE	30 + .5(73.1) + .4(84.5) + .1(87.0)	= 109.05		
KF	JF	50 + .3(74.8) + .2(89.2) + .5(117.0)	= 148.48	STAY IN K	112.33
	KF	30 + .3(73.1) + .2(84.5) + .5(87.0)	= 112.33		
KG	KG	30 + .2(73.1) + .4(84.5) + .4(87.0)	= 113.22	STAY IN K	113.22

TABLE 6

Period t = T- 5(Stages to go n =5)

Position Variable x_t	Decision Variables Allowed u_t	Expected Costs $z_n(x_t,u_t)$	Best Decision u_t^*	Corresponding Cost $f_n(x_t)=\min z_n(x_t,u_t)$
IE	IE	10 + .5(89.05) + .4(122.33) + .1(123.22)= 117.779	STAY IN I	117.779
	JE	40 + .5(104.78) + .4(118.48) + .1(143.22)= 154.104		
	KE	40 + .5(99.05) + .4(112.33) + .1(113.22)= 145.779		
IF	JF	40 + .3(104.78) + .2(118.48) + .5(143.22)= 166.740		
	KF	40 + .3(99.05) + .2(112.33) + .5(113.22)= 108.791	SWITCH TO K	108 791
IG	KG	40 + .2(99.05) + .4(112.33) + .4(113.22)= 110 030	SWITCH TO K	110.030
JE	IE	30 + .5(89.05) + .4(122.33) + .1(123.22)= 137.779		
	JE	20 + .5(104.78) + .4(118.48) + .1(143.22)= 134.104	STAY IN J	134.104
	KE	60 + .5(99.05) + .4(112.33) + .1(113.22)= 165.779		
JF	JF	20 + .3(104.78) + .2(118.48) + .5(143.22)= 146.740	STAY IN J	146.740
	KF	60 + .3(99.05) + .2(112.33) + .5(113.22)= 168.791		
JG	KG	60 + .2(99.05) + .4(112.33) + .4(113.22)= 170 030	SWITCH TO K	170 030
KE	IE	20 + .5(89.05) + .4(122.33) + .1(123.22)= 127.779	SWITCH TO I	127.779
	JE	50 + .5(104.78) + .4(118.48) + .1(143.22)= 164.104		
	KE	30 + .5(99.05) + .4(112.33) + .1(113.22)= 135.779		
KF	JF	50 + .3(104.78) + .2(118.48) + .5(143.22)= 176.740	STAY IN K	138.791
	KF	30 + .3(99.05) + .2(112.33) + .5(113.22)= 138.791		
KG	KG	30 + .2(99.05) + .4(112.33) + .4(113.22)= 140.030	STAY IN K	140.030

TABLE 7

Period t = T-6 (Stages to go n = 6)

Position Variable x_t	Decision Variables Allowed u_t	Expected Costs $z_n(x_t,u_t)$	Best Decision u_t^*	Corresponding Cost $f_n(x_t)=\min z_n(x_t,u_t)$
IE	IE	10 + .5(117.779) + .4(108.791) + .1(110.030) = 123.4	STAY IN K	123.4
	JE	40 + .5(134.104) + .4(146.740) + .1(170.030) = 182.7		
	KE	40 + .5(127.104) + .4(138.791) + .1(140.030) = 153.4		
IF	JF	40 + .3(134.104) + .2(146.740) + .5(170.030) = 194.6	SWITCH TO K	176.1
	KF	40 + .3(127.779) + .2(138.791) + .5(140.030) = 176.1		
IG	KG	40 + .2(127.779) + .4(146.740) + .4(140.030) = 180.2	SWITCH TO K	180.2
JE	IE	30 + .5( ) + .4( ) + .1( ) = 143.4	SWITCH TO I (first change)	143.4
	JE	20 + .5( ) + .4( ) + .1( ) = 162.4		
	KE	60 + .5( ) + .4( ) + .1( ) = 173.4		
JF	JF	20 + .3( ) + .2( ) + .5( ) = 174.6	STAY IN J	174.6
	KF	60 + .3( ) + .2( ) + .5( ) = 196.1	(and ultimately change into K)	
JG	KG	60 + .2( ) + .4( ) + .4( ) = 200.2	SWITCH TO K	200.2
KE	IE	20 + .5( ) + .4( ) + .1( ) = 133.4	SWITCH TO I	133.4
	JE	50 + .5( ) + .4( ) + .1( ) = 192.4		
	KE	30 + .5( ) + .4( ) + .1( ) = 153.4		
KF	JF	50 + .3( ) + .2( ) + .5( ) = 204.6	STAY IN K	166.1
	KF	30 + .3( ) + .2( ) + .5( ) = 166.1		
KG	KG	30 + .2( ) + .4( ) + .4( ) = 170.2	STAY IN K	170.2

Though we have not pushed the computation further, it is obvious from the figures already obtained that only one more change is liable to occur, namely, to switch from J to K when the weather is F, which puts J completely out of the optimal strategy.

After that, we will have reached a steady-state. This means that if n is the number of the iteration when this last change occurs, the subsequent strategy will always be the same for all processes involving more than these n steps.

This strategy can be stated as follows:

a) It is always recommended not to switch if the weather allows it, except for J which should be avoided systematically.

b) If one has to switch upwards or downwards the switch should be of the largest span allowed by the weather conditions. That is, for example, if one is in I and weather conditions are of type F, switching should be done to K and not to J. Conversely, if one is in K and can switch upwards, it is not worth switching to J, but a switch to I, if allowed, would be profitable.

Obviously, these conclusions depend strongly on the figures used. For instance, it is clear that, if switching to the intermediate technology J is never recommended,

this results from the fact that our hypothetical data show that first, it is relatively more costly to switch to J than to switch to or stay in I or K, and second, that the weather conditions F (during which the dirtiest allowed technology is J) are relatively unstable and therefore it is more profitable to stay in K, wait until weather conditions change to E or to G (which is more likely than to see F remain) and then switch to I or stay in K.

This very simple example shows already that the best strategy does not simply consist of adjusting the technology used to the dirtiest level allowed by weather conditions. However, this conclusion depends on the relative importance of switching and operating costs (which is arbitrary in our numerical example). Moreover, this relative weight varies as the process involves more and more stages : the expected value of the remaining stages increases and therefore reduces the relative importance of the stage under study (and especially, reduces within this stage the relative difference in costs between switching or not switching).

The above computations have been done by hand and without using any discount rate. It is very likely that the introduction of a discount rate would change the conclusions significantly. These are also sensitive to the structure of the transition matrix as well as to that of

the cost matrix. It would be interesting to study the asymptotic behaviour of the system by using real weather data. Unfortunately, the weather parameters which are relevant for air pollution control are not classified in clear types and are not generally presented under the form of a matrix of transition of types from one day to the next. Basic data processing in order to get this information in a chosen area is, however, possible and seems worthwhile.

Another difficulty consists in the estimation of the operating and switching costs associated with various technologies. For instance, let us assume that we consider an electric power generating plant equipped with a control device, say a desulphurizing device. Besides the cost of purchasing this piece of equipment, there is also a cost associated with its use. It may be profitable, if the weather allows it, not to operate this control device some days, and to discharge the gaseous effluent directly into the atmosphere. The degree of cleaning of the effluents may also vary and so may the associated costs. These points have been discussed extensively in our paper on the background analysis of the problem at hand.

We should therefore look for the costs of reducing the pollutants by 10%, 50%, or 100% by using the same device. One other variable, already stated at the beginning of this paper, is the use of various types of energy

sources: coal with high or low sulphur content, fuel, natural gas, etc.

The costs of the abatement strategy developed in the above example (6 stages process) can be derived as follows. If no abatement regulations apply, the plant would have used I over all the 6 periods, and the cost would have been (without discounting) equal to 60. With the existence of the abatement strategy, it is equal to 123 if in the initial state the weather conditions are E and the technology used is I. This strategy is more economical than using the cleanest technology on a permanent basis -- which would lead to a cost of 180 -- in spite of the high switching costs adopted in our example.

Chapter 4. Inclusion of Interperiod Trade-Offs in the Model

Let us recall that this problem, as stated in detail in Section 1, can be described as follows. At any stage, any technology I,J,K could be used with any weather condition E,F,G, provided the total concentration of pollutants resulting from this stage and all the preceding stages does not exceed an upper limit C.

Apparently, the algorithm developed for the special case cannot be applied here. Indeed, this algorithm proceeds from the last stage backwards. At each stage, it considers only the stages to go, adding one at each iter-

ation. This is made possible because the set of decisions available at each stage does not depend on the states or decisions of the previous stages; this set {U}, actually is the same for all periods.

In the general case, the set of decisions depends on the level of concentration reached as a result of the operations during the previous stages. Another approach seems then to be required.

However, we shall show that this problem can be reduced to the previous one by adding a third dimension to the state variable, namely the concentration of pollutants in the atmosphere at the beginning of each time period. By enumerating the possible combinations of these three dimensions, we can again identify a restricted set of decisions which is the same for each time period.

The algorithm developed above can then be applied to solve the general problem of this section, as indicated in the following paragraph.

4.1 Expression of the Problem in the Dynamic Programming Form

In order to keep the discrete character of all the problems examined here, we shall assume that the values that can be taken by the concentration of pollutants are discrete and that their number is finite. This can always be done by dividing the scale of concentrations into intervals and representing each interval by its mid-range value.

As each state of the system can be associated with each of the values of the finite set of concentrations, it is possible to consider concentration as an additional component of the state vector of the system the initial two components of which are the "technology" used and the weather condition.

The constraint that we assume to be imposed by regulation, namely that the concentration should not exceed a limit C_1 allows the definition of the possible values of the third component of the state variables. For instance, if $C^1 = 5$ and we consider only integer values of the concentrations, the possible values that the third component can assume are:

$$C^0=0 \quad C^1=1 \quad C^2=2 \quad C^3=3 \quad C^4=4 \quad C^5=5$$

A state of the system at time period t is then characterized by a vector x_t such as $x_t = [IEC^5]$ for instance. If the components associated with technology, weather and pollution concentration are respectively referred to as x^1, x^2, x^3, then a state vector is defined by $x_t = [x^1x^2x^3]$ (these are represented as row vectors, as opposed to column vectors, for writing convenience).

The above mentioned constraint allows also to define the set of possible decisions. Indeed, we assumed in Section 1 that a diffusion model provides us without uncertainty with the pollution concentration level added during a time period by any combination of technology and weather. For instance, in the example developed above, we may have the following correspondence:

Combinations of Technology and Weather	Units Added (ΔC)
IE	0
IF	1
IG	2
JE	-1
JF	0
JG	1
KE	-2
KF	-1
KG	0

Therefore, with the states of the system where the concentration component is equal to 5, it is not possible to make the following decisions as their effect is to add

At least one unit to the concentration, which is not allowed:

{IF, IG, JG}

With the states of the system where the concentration component is 4, we must exclude IG as a possible decision as it adds more than one unit to the concentration.

For all other state variables, a decision to use any state of the system can be made.

This can be formalized by stating that all states such as $x^3 + \Delta C(x^1,x^2) > C^1$ where $\Delta C(x^1,x^2)$ is the concentration added, when the system operates with technology x^1 and weather conditions x^2 must be eliminated from the set of possible decisons, for all time periods.

By doing so, we define a restricted set of decision variables which is independent of time. This make it possible to use the algorithm developed in the preceding section. The only difference is that the dimensions of the problem are larger.

4.2 Application to the Example Developed in the Preceding Section

The sets of position and decision varialbes are indicated in the figure below. The column referred to as "weather" indicates the set of possible states of the system after the random change due to weather has occurred.

END OF PERIOD T-2	BEGINNING OF PERIOD T-1	END OF PERIOD T-1	BEGINNING OF PERIOD T
IEC_1	IEC_1	IEC_1	IEC_1
IEC_2	IEC_2	IEC_2	IEC_2
IEC_3	IEC_3	IEC_3	IEC_3
IFC_1	IFC_1	IFC_1	IFC_1
IFC_2	IFC_2	IFC_2	IFC_2
IFC_3		IFC_3	
IGC_1	IGC_1	IGC_1	IGC_1
IGC_2		IGC_2	
IGC_3		IGC_3	
JEC_1	JEC_1	JEC_1	JEC_1
JEC_2	JEC_2	JEC_2	JEC_2
JEC_3	JEC_3	JEC_3	JEC_3
JFC_1	JFC_1	JFC_1	JFC_1
JFC_2	JFC_2	JFC_2	JFC_2
JFC_3	JFC_3	JFC_3	JFC_3
JGC_1	JGC_1	JGC_1	JGC_1
JGC_2	JGC_2	JGC_2	JGC_2
JGC_3		JGC_3	
KEC_1	KEC_1	KEC_1	KEC_1
KEC_2	KEC_2	KEC_2	KEC_2
KEC_3	KEC_3	KEC_3	KEC_3
KFC_1	KFC_1	KFC_1	KFC_1
KFC_2	KFC_2	KFC_2	KFC_2
KFC_3	KFC_3	KFC_3	KFC_3
KGC_1	KGC_1	KGC_1	KGC_1
KGC_2	KGC_2	KGC_2	KGC_2
KGC_3	KGC_3	KGC_3	KGC_3
CHANGE IN THE WEATHER COMPONENT ONLY (RANDOM)	CHANGE IN THE ABATEMENT LEVEL COMPONENT ONLY (DECISION + COST)	CHANGE IN THE POLLUTION CONCENTRATION ONLY (DETERMINISTIC)	CHANGE IN THE WEATHER COMPONENT ONLY (RANDOM)	CHANGE IN THE ABATEMENT LEVEL COMPONENT ONLY (DECISION + COST)

For instance, during period T-2, the system has operated under state [JEC_2], A random change in weather switches the system to state [JFC_2] (only the weather component is changed). At the beginning of time period T-1, the system is in this last state. A decision is made to switch it to [IFC_2] (only the technology component is changed). According to our diffusion model, operating with I when the weather is E during a complete time period adds one unit to the concentration. Therefore, at the end of time period T-1, the system moves to [IFC_3] (only the concentration component has changed). At this time a weather change occurs again, etc.

Knowing the matrices of weather transitions and of switching and total costs and assuming that the time horizon of the problem is the end of period T-1, the algorithms developed in the previous section can be applied without modification.

The description above assumes that the addition of pollution is not associated with a cost. This is only an apparance. Indeed, when a cost is associated with the state variable's third dimension, it will simply show as an additional cost within the cost vector $L(x_t)$. Such a cost can represent either a tax that the firm has to pay, or an estimation of the cost occuring to the community, or both.

Chapter 5: Proof of Existence and Unicity of the Optimal Solution by the Algorithm

5.1 Alternative Formulation of the Problem

The system studied is characterized by the same three-dimensional state variables x_t as above. This state is observed at the beginning of period t. Then a decision u_t is made. This is an input to the system which results in transferring the state from x_t to $\bar{x}_t$, the state of the system after the decision. There are N state variables and m decision variables. We assume here that the set of decision variables is the same during all periods. The way this can be done in connection with the present problem which has been described in the preceding chapter.

After a decision is made and during all the time periods, the only change in the state of the system occurs in the pollution concentration component.

At the end of the time period, what is observed -- that is, the system's output -- is the state of the system in terms of technology, weather and concentration. These have been so far assumed observable with certainty. We shall continue to make this assumption in the present paragraph.

Between the end of one period and the beginning of the next one, the system evolves at random, according to a Markov process. The entries of the transition matrix ϕ_t of this process will be written:

$$\phi_t(l,i) = \text{Prob}\,[x_{t+1} = i | u_t = l]$$

We shall assume that the intiial states are either known or their probabilities are known and that none of these probabilities is zero. In that case, the N-vector of state probabilities is known at the beginning of every time period.

Let $\psi_t(i) = \text{Prob}\,[x_t = i]$ for $i = 1,\ldots,N$

where $\sum_{i=1}^{N} \psi_t(i) = 1$ and $\psi_t(i) \geq 0$ $\quad i,t$

Note that $\psi_t(i) = \phi^t . \psi_o(i)$.

A set of difference equations will be obtained for describing the transition of vector ψ_t in terms of the input u_t as follows.

Consider the matrix D_t whose entries $D_t(k,l)$ are defined by:

$$D_t(k,l) = \text{Prob}\,[u_t = l \mid x_t = k] \quad \text{for } k = 1,\ldots,N$$

with $0 \leq D_t(k,l) \leq 1$ $\quad \forall\ k,l,t$

$$\sum_{l=1}^{m} D_t(k,l) = 1 \quad \forall\ k,l$$

The variables $D_t(k,l)$ will play the role of controls. The matrix D_t is called a decision rule or a control rule or a strategy.

It can be easily seen that:

$$\psi_{t+1}(i) = \sum_{k=1}^{N} \sum_{l=1}^{m} \text{Prob}\,[x_{t+1} = i \mid x_t = k,\ u_t = l]$$

$$= \sum_{k=1}^{N} \sum_{l=1}^{m} \psi_t(k)\ D_t(k,l)\ \phi_t(l,i)$$

The value of u_t that has to be used at each time period t is sought so as to minimize the following criterion function:

$$J = E[\sum_{t=0}^{T-1} L_t(\bar{x}_t) + M_t(x_t,\bar{x}_t)]$$

where E denotes the expectation, $M_t(x_t,\bar{x}_t)$ denotes the cost of switching from state x_t to state $\bar{x}_t$ and $L_t(\bar{x}_t)$ the operating cost for state $\bar{x}_t$ (this cost can also include taxes for polluting and the damages due to pollution incurred by the community).

The subscript t associated with the costs refers to the fact that costs might not be the same for each period of time or simply that they may be discounted by a function of time.

This function can be generalized first by introducing a cost for the terminal state. This terminal charge can add flexibility by allowing a global accounting of what happens after the last period.

In the most general case, the criterion becomes:

$$J = E[L_{T-1}(x_t) + \sum_{t=1}^{T-2} L_t(\bar{x}_t) + M_t(x_t,\bar{x}_t)]$$

We can rewrite the criterion in terms of ψ_t, D_t as follows:

$$J = \sum_{t=0}^{T-2} \sum_{l=1}^{m} \sum_{k=1}^{N} \psi_t(k).D_t(k,l)\, M_t(k,l)$$

$$(+ \sum_{k=1}^{N} \psi_{t-1}(k).L_{t-1}(k))$$

$$(+ \sum_{t=0}^{T-2} \sum_{k=1}^{N} \psi_t(k).L_t(k))$$

The problem is then to maximize J with respect to u

with: $\phi_t(l,i) = \text{Prob}\,[x_{t+1} = i | u_t = l]$

$D_t(k,l) = \text{Prob}\,[u_t = l | x_t = k$

$\psi_t(i) = \text{Prob}\,[x_t = i]$

$$\psi_{t+1}(i) = \sum_{l=1}^{m} \sum_{K=1}^{N} \psi_t(k).D_t(k,l).\phi_t(l,i)$$

and where the constraints are the following:

$$0 \leq D_t(k,l) \leq 1 \qquad k,l,t$$

$$\sum_{l=1}^{m} D_t(k,l) = 1 \qquad l,t$$

$$\psi_{t=1} \geq 0 \qquad \text{and} \qquad \sum_{i=1}^{N} \psi_{t=1}(i) = 1$$

5.2 Solution of the Problem

The solution to this problem is based on two theorems: the first of these theorems, proved by Derman[1], establishes that when the final state is observable without uncertainty, then the optimal $D_t(k,l)$ are unique and all equal to zero or 1; in otherwords, the strategy is pure.

---

[1] C. Derman, Finite State Markovian Decision Process, Academic Press, 1970.

The second theorem is an extension of the Pontryagin maximum principle to the case where state variables are not continuous. It is known as the Discrete Maximum principle and has been developed by several authors, in particular Fan and Wang[2]. This theorem establishes that the minimum of J is obtained when the Hamiltonian of J is stationary along the optimal trajectory. In order to understand this theorem, let us recall the definition of the Hamiltonian of a function

$$J = L[x(T)] + \sum_{t=1}^{T-1} M_t(x_t, u_t)$$

where $x_{t+1} = f_t(x_t, u_t, t)$, x_o given, $t=1,\ldots,T$

The Hamiltonian is by definition $H \overset{\Delta}{=} \lambda'_{t+1} f_t + M_t$

where $\lambda_t = \frac{\partial H}{\partial x_t}$ (or $\lambda_{t+1} = \frac{\partial H}{\partial f_t}$) the λ_t are called the "adjoint" variables

$$\lambda_T = \frac{\partial L[x(T)]}{\partial x(T)}$$

In our particular case:

$$H = \sum_{i=1}^{N} \sum_{l=1}^{m} \sum_{k=1}^{N} \lambda_{t+1}(i)\ \psi_t(k).D_t(k,l).\phi_t(l,i)$$

$$+ \sum_{l=1}^{m} \sum_{k=1}^{N} \psi_t(k).D_t(k,l).M_t(k,l)$$

$$(+ \sum_{k=1}^{N} \psi_t(k).L_t(k))$$

---

[2] L. Fan and C. Wang, The Discrete Maximum Principle, John Wiley, 1964.

where l_o is found by the relation $a_t(l_o,k) = \min_l a_t(l,k)$

Notice that: $D_t(k,l)$ depends explicitly on λ_{t+1} and not on ϕ_t and that the adjoint equations for λ_t are independent of ϕ_t

Therefore, the optimal policy $D_t(o)$ is given directly by solving only the set of N adjoint difference equations $\lambda_t = \frac{\partial H}{\mu \psi_t}$ with final condition $\lambda_T = L_T$ and the expressions defining l_o.

It should be noted that the optimal control $D_t(o)$ is independent of ψ_t for all t.

For these problems, the Lagrange multipliers have an elegant interpretation: $\lambda_t(i)$ represents the optimal value of the expected cost of a system which starts at time t in ith state and terminates at time T-1 (to be proved).

5.3 Numerical Example (The following example is independent of the one developed in the previous section.)

Consider a system with 4 states A,B,C,D and 3 inputs A,B,C. The transition probability matrices and the cost of transitions are given in the tables below. It is desired to find the strategy matrices D_1 and D_2 to minimize J

$$J = \sum_{t=1}^{2} \sum_{l=1}^{3} \sum_{k=1}^{4} \psi_t(k).D_t(k,l).M_t(k,l) + \sum_{k=1}^{4} L_T(k).\psi_t(k)$$

$$\psi_{t+1}(t) = \sum_{l=1}^{3} \sum_{k=1}^{4} \psi_t(k).D_t(k,l)\phi(l,i)$$

$\psi_t(l,i)$ =

	A	B	C	D
A	.5	.3	.1	.1
B	.4	.2	.2	.2
C	.1	.3	.3	.3

$M_t(k,l)$ =

	A	B	C
A	10	20	30
B	20	10	50
C	10	10	10
D	30	30	20

$L_T(i)$ =

A	1
B	2
C	3
D	4

$D_t(k,l)$ =

	A	B	C
A	.	.	.
B	.	.	.
C	.	.	.
D	.	.	.

$\psi_t(k)$ =

A	.
B	.
C	.
D	.

$\lambda_{t+1}(i)$ =

A	.
B	.
C	.
D	.

$$H = \sum_i \lambda_{t+1}(i).\psi_{t+1}(i) + \sum_l \sum_k \psi_t(k).D_t(k,l).M(k,l)$$

$$\lambda_t(k) = \sum_{l=1}^{M} D_t(k,l).\ a_t(l,k)$$

$$a_t(l,k) = \sum_i \lambda_{t+1}(i).\phi_t(l,i) + M_t(k,l)$$

$$\lambda_T(k) = L_T(k)$$

The minimum of H is obtained when $D_t(k,1)=1$ if $1=1_o$

$D_t(k,1) = 0$ otherwise

$$a_t(1_o,k) = \min_{1}[\ \Sigma\lambda_{t+1}(i)\ \phi_t(1,i) + M_t(k,1)]$$

Last (Third) Period: $\lambda_3 = L_3 = \begin{matrix}1\\2\\3\\4\end{matrix}$

Second Period:

As 1 can take only a discrete number of values, 1_o is obtained by giving 1 in $a_t(1,k)$ each of its values and selecting for 1_o the one that will minimize $a_t(1,k)$.

This is done for each value of k.

K = A 1 = A $a_2(A,A)= 1(.5)+2(.3)+3(.1)+4(.1)+10 = 11.8$

1 = B $a_2(B,A)= 1(.4)+2(.2)+3(.2)+4(.2)+20 = 22.2$

1 = c $a_2(C,A)= 1(.1)+2(.3)+3(.3)+4(.3)+30 = 32.8$

Therefore $a_2(A) = \begin{matrix}11.8\\22.2\\32.8\end{matrix}$ and the smallest value corresponds to 1_o = A.

The same calculations with K = B, K = C, K = D yield:

$a_2(B) = \begin{matrix}21.8\\12.2\\52.8\end{matrix}$ whose smallest value is 12.2, corresponding to 1_o = B.

$$a_2(C) = \begin{matrix} 11.8 \\ 12.2 \\ 12.8 \end{matrix}$$ whose smallest value is 11.8, corresponding to l_o = A

$$a_2(D) = \begin{matrix} 31.8 \\ 32.2 \\ 22.8 \end{matrix}$$ whose smallest value is 22.8, corresponding to l_o = C.

(In all these calculations, only the last term of each equation ($M_t(k,1)$ changes.)

The decision matrix for period 2 is therefore:

$$D_2 = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{matrix}$$

We now have to calculate the values of $\lambda_2(k)$ in order to proceed to the first period:

$$\lambda_2(k) = \sum_1 D_2(k,1)\ a_2(1,k) + L_t(k)$$

for K = A $\quad \lambda_2(A) = 1(11.8) + 0(22.2) + 0(32.8) + 1 = 12.8$

Similarly for:

K = B $\quad \lambda_2(B) = 14.2$

K = C $\quad \lambda_2(C) = 14.8$

K = D $\quad \lambda_2(D) = 26.8$

First Period

K=A 1=A $a_1(A,A)=12.8(.5)+14.2(.3)+14.8(.1)+26.8(.1)+10 = 24.82$

1=B $a_1(B,A)=12.8(.4)+14.2(.2)+14.8(.2)+26.8(.2)+20 = 36.68$

1=C $a_1(C,A)=12.8(.1)+14.2(.3)+14.8(.3)+26.8(.3)+30 = 48.02$

$a_1(A) = \begin{matrix}24.82\\36.68\\48.02\end{matrix}$ whose smallest value is 24.82 corresponding to $1_o = A$

Similarly:

$a_1(B) = \begin{matrix}34.82\\26.68\\68.02\end{matrix}$ whose smallest value is 26.68 corresponding to $1_o = B$

$a_1(C) = \begin{matrix}24.82\\26.68\\28.02\end{matrix}$ whose smallest value is 24.82 corresponding to $1_o = A$

$a_1(D) = \begin{matrix}44.82\\46.68\\38.02\end{matrix}$ whose smallest value is 36.02 corresponding to $1_o = C$

The decision matrix for period 1 is therefore $D_1 = \begin{bmatrix}1&0&0\\0&1&0\\1&0&0\\0&0&1\end{bmatrix}$

In this example, D_1 is equal to D_2. This is only caused by the figures which have been chosen arbitrarily. With some other values, we would not have this equality. However, if the process is pursued over a large number of periods, the matrices D_t will end up to be all equal and the

phenomenon will stabilize, as discussed in the first chapters of the present paper.

This example illustrates the exact similarity of the algorithms obtained by the two approaches and sheds light on the meaning of the Lagrange multipliers $\lambda_t(i)$. These represent, as announced above, the optimal value of the expected cost of a system which starts at time t in the i^{th} state and terminates at time T-1.